

# Optimal Control of a Parabolic Distributed Parameter System Using a Barycentric Shifted Gegenbauer Pseudospectral Method

Kareem T. Elgindy\*

*Mathematics Department, Faculty of Science, Assiut University, Assiut 71516, Egypt*

## SUMMARY

In this paper, we introduce a novel pseudospectral method for the numerical solution of optimal control problems governed by a parabolic distributed parameter system. The infinite-dimensional optimal control problem is reduced into a finite-dimensional nonlinear programming problem through shifted Gegenbauer quadratures constructed using a stable barycentric representation of Lagrange interpolating polynomials and explicit barycentric weights for the shifted Gegenbauer-Gauss (SGG) points. A rigorous error analysis of the method is presented, and a numerical test example is given to show the accuracy and efficiency of the proposed pseudospectral method.

**KEY WORDS:** Barycentric interpolation; Integration matrix; Optimal control; Pseudospectral method; Shifted Gegenbauer polynomial; Shifted Gegenbauer quadrature.

## 1. INTRODUCTION

Optimal control theory has attracted much attention since the 1950s after the arrival of digital computers, which provided the impetus for the applications of the branch to many complicated problems; cf. [Elgindy (2013)]. One of the primary objectives of this significant branch is to find the control signals that will cause a process to satisfy certain physical constraints while optimizing some performance criterion. Analytical methods can solve only fairly simple problems, therefore much research in this area has been devoted to developing accurate and efficient numerical methods to obtain approximate solutions instead of looking for closed form exact solutions that could be very cumbersome or either impossible to determine.

In this paper, we present a novel and powerful numerical method for the solution of an optimal control problem governed by a parabolic distributed parameter system, which has been recently solved numerically by [Rad et al. (2014)] using radial basis functions. The present method belongs to the class of pseudospectral methods that were largely developed in the 1970s for solving partial differential equations (PDEs), and impetuously imposed itself strongly as ‘one of the big three technologies for the numerical solution of PDEs’ [Trefethen (2000)]. The proposed pseudospectral method is a strong tool that exhibits exponential convergence rates, and able to produce accurate approximations using a relatively very small number of collocation points. The central idea in this work is to exploit the well-conditioning of numerical integration operators via recasting the optimal control problem into its integral form. We then approximate the involved integral operators by integration matrices based on shifted Gegenbauer quadratures that can be constructed efficiently using the recently developed Gegenbauer quadratures of [Elgindy (2016b)]. The novel quadratures are defined based on the stable barycentric representation of Lagrange interpolating polynomials and the explicit barycentric weights for the shifted Gegenbauer-Gauss (SGG) points. The pseudospectral method eventually endeavors to reduce the infinite-dimensional optimal control problem to a finite-dimensional nonlinear programming problem with linear constraints that can be solved easily using standard numerical optimization solvers.

---

\*Correspondence to: Mathematics Department, Faculty of Science, Assiut University, Assiut 71516, Egypt

The remainder of this paper is structured as follows: In Sections 2 and 3, we state the mathematical formulation of the optimal control problem and its integral formulation, respectively. In Section 4, we present the novel barycentric shifted Gegenbauer pseudospectral method (BSGPM) for discretizing the integral optimal control problem. Section 5 is devoted for a rigorous error and convergence analysis of the proposed method to verify the spectral decay of the error for increasing number of collocation points. A numerical test example is presented in Section 6 to assess the accuracy and efficiency of the proposed method followed by some concluding remarks in Section 7.

## 2. PROBLEM STATEMENT

In this study, we are interested in finding the control function  $u : D_{L,t_f}^2 \rightarrow \mathbb{R}$ , and the corresponding state function  $x : D_{L,t_f}^2 \rightarrow \mathbb{R}$ , that minimize the quadratic cost functional,

$$J(x, u) = \int_0^{t_f} \int_0^L (r_1 x^2(y, t) + r_2 u^2(y, t)) dy dt, \quad (2.1)$$

subject to the one-dimensional diffusion equation,

$$x_t(y, t) = x_{yy}(y, t) + u(y, t), \quad (2.2)$$

with the initial condition,

$$x(y, 0) = f(y), \quad 0 \leq y \leq L, \quad (2.3)$$

and the boundary conditions,

$$x_y(0, t) = 0, \quad 0 \leq t \leq t_f, \quad (2.4)$$

$$x_y(L, t) = 0, \quad 0 \leq t \leq t_f, \quad (2.5)$$

where  $D_{L,t_f}^2 = [0, L] \times [0, t_f]$ , and  $L, t_f, r_1, r_2 \in \mathbb{R}^+$ .

## 3. THE INTEGRAL FORMULATION OF THE OPTIMAL CONTROL PROBLEM

Let,

$$I_{q,\tilde{y}}^{(y)}(\vartheta(y, t)) = \int_0^{\tilde{y}} \int_0^{\sigma_{q-1}} \dots \int_0^{\sigma_2} \int_0^{\sigma_1} \vartheta(\sigma_0, t) d\sigma_0 d\sigma_1 \dots d\sigma_{q-2} d\sigma_{q-1}, \quad (3.1)$$

$$I_{q,\tilde{t}}^{(t)}(\vartheta(y, t)) = \int_0^{\tilde{t}} \int_0^{\sigma_{q-1}} \dots \int_0^{\sigma_2} \int_0^{\sigma_1} \vartheta(y, \sigma_0) d\sigma_0 d\sigma_1 \dots d\sigma_{q-2} d\sigma_{q-1}, \quad (3.2)$$

denote the  $q$ -fold integrals of any integrable bivariate function  $\vartheta(y, t)$  w.r.t.  $y$  and  $t$ , respectively, for any positive real numbers  $\tilde{y} \in [0, L]$  and  $\tilde{t} \in [0, t_f]$ . Using the substitution,

$$x_{yy}(y, t) = \phi(y, t), \quad (3.3)$$

for some unknown function  $\phi$ , we can recover the unknown state function  $x$  and its first-order partial derivative  $x_y$  in terms of  $\phi$  via successive integration and the boundary condition (2.4) as follows:

$$x_y(y, t) = I_{1,y}^{(y)}(\phi(y, t)), \quad (3.4)$$

$$x(y, t) = I_{2,y}^{(y)}(\phi(y, t)) + c_1(t), \quad (3.5)$$

where  $c_1(t)$  is some arbitrary function in  $t$ . The boundary condition (2.5) yields,

$$I_{1,L}^{(y)}(\phi(y, t)) = 0. \quad (3.6)$$

Now integrating the Dynamical System Eq. (2.2) with respect to  $t$  and using the initial condition (2.3) gives,

$$x(y, t) = I_{1,t}^{(t)}(\phi(y, t) + u(y, t)) + f(y). \quad (3.7)$$

Equating Eqs. (3.5) and (3.7) yields,

$$I_{2,y}^{(y)}(\phi(y, t)) + c_1(t) = I_{1,t}^{(t)}(\phi(y, t) + u(y, t)) + f(y). \quad (3.8)$$

Since Eq. (3.8) is satisfied for all  $(y, t) \in D_{L,t_f}^2$ ,  $c_1(t)$  can be determined by setting  $y = 0$ ; so

$$c_1(t) = I_{1,t}^{(t)}(\phi(0, t) + u(0, t)) + f(0), \quad (3.9)$$

and the integral one-dimensional diffusion equation can be written as,

$$I_{2,y}^{(y)}(\phi(y, t)) + I_{1,t}^{(t)}(\phi(0, t) - \phi(y, t) + u(0, t) - u(y, t)) = f(y) - f(0). \quad (3.10)$$

The cost functional can also be written as,

$$J(\phi, u) = I_{1,t_f}^{(t)} I_{1,L}^{(y)} \left( r_1 \left( I_{1,t}^{(t)}(\phi(y, t) + u(y, t)) + f(y) \right)^2 + r_2 u^2(y, t) \right). \quad (3.11)$$

Hence, the integral optimal control problem is to find the control function  $u : D_{L,t_f}^2 \rightarrow \mathbb{R}$ , and the corresponding second-order derivative of the state function  $\phi : D_{L,t_f}^2 \rightarrow \mathbb{R}$ , w.r.t.  $y$ , that minimize the quadratic cost functional (3.11) subject to Eqs. (3.10) and (3.6).

#### 4. THE BSGPM

Let  $\mathbb{S}_{l,n}^{(\alpha)} = \{x_{l,n,k}^{(\alpha)}, k = 0, \dots, n\}$ , denote the set of the zeroes (SGG nodes) of the  $(n+1)$ th-degree shifted Gegenbauer polynomial,  $G_{l,n+1}^{(\alpha)}(x)$ , defined on the interval  $[0, l]$ , for any  $l \in \mathbb{R}^+, n \in \mathbb{Z}^+, \alpha > -1/2$ , and let  $\varpi_{l,n,i}^{(\alpha)}, i = 0, \dots, n$ , be their corresponding Christoffel numbers. We lay a grid of SGG nodes,  $\left(y_{L,N_y,i}^{(\alpha)}, t_{t_f,N_t,j}^{(\alpha)}\right), i = 0, \dots, N_y; j = 0, \dots, N_t$ , on the rectangular domain  $D_{L,t_f}^2$ , for some  $N_y, N_t \in \mathbb{Z}^+$ , sorted ascendingly as  $y_{L,N_y,N_y+1}^{(\alpha)} = 0 < y_{L,N_y,0}^{(\alpha)} < y_{L,N_y,1}^{(\alpha)} < \dots < y_{L,N_y,N_y}^{(\alpha)} < y_{L,N_y,N_y+2}^{(\alpha)} = L; t_{t_f,N_t,N_t+1}^{(\alpha)} = 0 < t_{t_f,N_t,0}^{(\alpha)} < t_{t_f,N_t,1}^{(\alpha)} < \dots < t_{t_f,N_t,N_t}^{(\alpha)} < t_{t_f,N_t,N_t+2}^{(\alpha)} = t_f$ , and approximate the function  $\phi$  by interpolation at the internal Gauss nodes. For simplicity, let us denote  $\phi\left(y_{L,N_y,s}^{(\alpha)}, t_{t_f,N_t,k}^{(\alpha)}\right)$ , by  $\phi_{s,k} \forall s, k$ . The polynomial interpolant of  $\phi$  in two dimensions can be written in Lagrange form as follows [Elgindy (2016a)]:

$$P_{N_y,N_t}\phi(y, t) = \sum_{s=0}^{N_y} \sum_{k=0}^{N_t} \phi_{s,k} \mathcal{L}_{N_y,N_t,s,k}^{(\alpha)}(y, t), \quad (4.1)$$

where  $\mathcal{L}_{N_y,N_t,s,k}^{(\alpha)}(y, t)$ ,  $s = 0, \dots, N_y; k = 0, \dots, N_t$ , are the bivariate Lagrange interpolating polynomials defined by,

$$\mathcal{L}_{N_y,N_t,s,k}^{(\alpha)}(y, t) = \mathcal{L}_{B,N_y,s}^{(\alpha)}(y) \mathcal{L}_{B,N_t,k}^{(\alpha)}(t) \forall s, k, \quad (4.2)$$

$$\mathcal{L}_{B,N_y,s}^{(\alpha)}(y) = \frac{\xi_{L,N_y,s}^{(\alpha)}}{y - y_{L,N_y,s}^{(\alpha)}} / \sum_{j=0}^{N_y} \frac{\xi_{L,N_y,j}^{(\alpha)}}{y - y_{L,N_y,j}^{(\alpha)}} \forall s, \quad (4.3)$$

$$\mathcal{L}_{B,N_t,k}^{(\alpha)}(t) = \frac{\xi_{t_f,N_t,k}^{(\alpha)}}{t - t_{t_f,N_t,k}^{(\alpha)}} / \sum_{j=0}^{N_t} \frac{\xi_{t_f,N_t,j}^{(\alpha)}}{t - t_{t_f,N_t,j}^{(\alpha)}} \forall k, \quad (4.4)$$

and  $\xi_{l,n,i}^{(\alpha)}$ ,  $i = 0, \dots, n$ , are the shifted barycentric weights defined by,

$$\xi_{l,n,i}^{(\alpha)} = 2(-1)^i \sqrt{4^\alpha l^{-2(1+\alpha)} \left(l - x_{l,n,i}^{(\alpha)}\right) x_{l,n,i}^{(\alpha)} \varpi_{l,n,i}^{(\alpha)}} \quad \forall l, n; i = 0, \dots, n. \quad (4.5)$$

Similarly, we can define the polynomial interpolant of  $u$  in two dimensions as follows:

$$P_{N_y, N_t} u(y, t) = \sum_{s=0}^{N_y} \sum_{k=0}^{N_t} u_{s,k} \mathcal{L}_{N_y, N_t, s, k}^{(\alpha)}(y, t), \quad (4.6)$$

The barycentric Lagrange interpolation enjoys several advantages such as being scale-invariant and forward stable for Gauss sets of interpolating points, which makes it very efficient in practice; cf. [Elgindy (2016b)].

Let  $\mathbf{P}_{B,n}^{(q)} = \left(p_{B,n,i,j}^{(q)}\right)$ ,  $i, j = 0, \dots, n$ , be the  $q$ th-order barycentric Gegenbauer integration matrix, for some  $q \in \mathbb{Z}^+$ , as defined by [Elgindy (2016b)], and denote its  $i$ th-row vector  $\left[l p_{B,n,i,0}^{(q)}, l p_{B,n,i,1}^{(q)}, \dots, l p_{B,n,i,n}^{(q)}\right]$  by  ${}_l \mathbf{P}_{B,n,i}^{(q)}$ . Similar to [Elgindy (2016a), Eqs. (4.42)], we can generate the  $q$ th-order barycentric shifted Gegenbauer integration matrix (BSGIM),  ${}_l \mathbf{P}_{B,n}^{(q)} = \left(l p_{B,n,i,j}^{(q)}\right)$ ,  $i, j = 0, \dots, n$ , through the following useful relation:

$${}_l \mathbf{P}_{B,n}^{(q)} = \left(\frac{l}{2}\right)^q \mathbf{P}_{B,n}^{(q)}. \quad (4.7)$$

By construction, we find that

$$I_{1, y_{L, N_y, i}}^{(y)} P_{N_y, N_t} \phi(y, t_{f, N_t, j}^{(\alpha)}) = \sum_{s=0}^{N_y} L p_{B, N_y, i, s}^{(1)} \phi_{s, j}, \quad (4.8a)$$

$$I_{1, t_{f, N_t, j}}^{(t)} P_{N_y, N_t} \phi(y_{L, N_y, i}^{(\alpha)}, t) = \sum_{k=0}^{N_t} t_f p_{B, N_t, j, k}^{(1)} \phi_{i, k}, \quad (4.8b)$$

for each  $i = 0, \dots, N_y$ ;  $j = 0, \dots, N_t$ , which is also true for  $P_{N_y, N_t} u$ . Let  $\mathbf{1} \in \mathbb{R}^{N_y+1}$ , be the all-ones vector, and “ $\otimes$ ” denotes the Kronecker product. The discrete integral one-dimensional diffusion equation can be written at the SGG mesh grid as,

$${}_L \mathbf{P}_{B, N_y, i}^{(2)} \left( \phi \left( \mathbf{y}_{L, N_y}^{(\alpha)}, t_{f, N_t, j}^{(\alpha)} \right) \right) + t_f \mathbf{P}_{B, N_t, j}^{(1)} \left( \phi \left( \mathbf{y}_{L, N_y, N_y+1}^{(\alpha)}, \mathbf{t}_{f, N_t}^{(\alpha)} \right) - \phi \left( \mathbf{y}_{L, N_y, i}^{(\alpha)}, \mathbf{t}_{f, N_t}^{(\alpha)} \right) + u \left( \mathbf{y}_{L, N_y, N_y+1}^{(\alpha)}, \mathbf{t}_{f, N_t}^{(\alpha)} \right) - u \left( \mathbf{y}_{L, N_y, i}^{(\alpha)}, \mathbf{t}_{f, N_t}^{(\alpha)} \right) \right) = \mathbf{F}, \quad (4.9)$$

for  $i = 0, \dots, N_y$ ;  $j = 0, \dots, N_t$ , where  $\mathbf{y}_{L, N_y}^{(\alpha)} = \left[ y_{L, N_y, 0}^{(\alpha)}, y_{L, N_y, 1}^{(\alpha)}, \dots, y_{L, N_y, N_y}^{(\alpha)} \right]^T$ ,  $\mathbf{t}_{f, N_t}^{(\alpha)} = \left[ t_{f, N_t, 0}^{(\alpha)}, t_{f, N_t, 1}^{(\alpha)}, \dots, t_{f, N_t, N_t}^{(\alpha)} \right]^T$ ;

$$\mathbf{F} = f \left( \mathbf{y}_{L, N_y}^{(\alpha)} \right) - f \left( \mathbf{y}_{L, N_y, N_y+1}^{(\alpha)} \right) \otimes \mathbf{1}. \quad (4.10)$$

Let  $\mathbf{P}_{B, N_y, N_y+2}^{(1)} = \left[ p_{B, N_y, N_y+2, 0}^{(1)}, p_{B, N_y, N_y+2, 1}^{(1)}, \dots, p_{B, N_y, N_y+2, N_y}^{(1)} \right]$ , be the barycentric Gegenbauer integration vector required for approximating definite integrals over the interval  $[-1, 1]$  as described by [Elgindy (2016b), Algorithms 6 and 7]. Similar to Eq. (4.7), we can easily construct the shifted barycentric Gegenbauer integration vector,  ${}_L \mathbf{P}_{B, N_y, N_y+2}^{(1)} = \left[ L p_{B, N_y, N_y+2, 0}^{(1)}, L p_{B, N_y, N_y+2, 1}^{(1)}, \dots, L p_{B, N_y, N_y+2, N_y}^{(1)} \right]$ , by the following formula:

$${}_L \mathbf{P}_{B, N_y, N_y+2}^{(1)} = \frac{L}{2} \mathbf{P}_{B, N_y, N_y+2}^{(1)}. \quad (4.11)$$

Hence, the discrete form of the boundary condition (3.6) is given by,

$${}_L \mathbf{P}_{B, N_y, N_y+2}^{(1)} \phi \left( \mathbf{y}_{L, N_y}^{(\alpha)}, t_{f, N_t, j}^{(\alpha)} \right) = 0, \quad j = 0, \dots, N_t. \quad (4.12)$$

The discrete cost functional can also be written as,

$$J_{N_y, N_t} = \sum_{l=0}^{N_t} t_f p_{B, N_t, N_t+2, l}^{(1)} \sum_{k=0}^{N_y} L p_{B, N_y, N_y+2, k}^{(1)} (r_1 x_{k, l}^2 + r_2 u_{k, l}^2), \quad (4.13)$$

where,

$$u_{k, l} = u \left( y_{L, N_y, k}^{(\alpha)}, t_{t_f, N_t, l}^{(\alpha)} \right) \forall k, l, \quad (4.14)$$

$$x_{k, l} = x \left( y_{L, N_y, k}^{(\alpha)}, t_{t_f, N_t, l}^{(\alpha)} \right) \approx \sum_{s=0}^{N_t} t_f p_{B, N_t, l, s}^{(1)} (\phi_{k, s} + u_{k, s}) + f_k \forall k, l; \quad (4.15)$$

$$f_k = f \left( y_{L, N_y, k}^{(\alpha)} \right) \forall k. \quad (4.16)$$

Now to put the pointwise representation of the discrete dynamics (4.9) and constraints (4.12) into a standard matrix system form, we introduce the mapping  $n = \text{index}(i, j) : n = i + j(N_y + 2)$ , and set  $L' = N_y + N_t + N_y N_t$ ,  $L'' = N_y + N_t(N_y + 2)$ ,

$$(\phi_{i, j}) = (\hat{\phi}_n) = \hat{\phi} \in \mathbb{R}^{L''+2}; \quad (4.17)$$

$$(u_{i, j}) = (\hat{u}_n) = \hat{u} \in \mathbb{R}^{L''+2}. \quad (4.18)$$

We also define the elements of the auxiliary matrices  $\mathbf{A} \in \mathbb{R}^{(L'+1) \times (L''+2)}$  and  $\mathbf{B} \in \mathbb{R}^{(L'+1) \times (L''+2)}$ , and the column vector  $\tilde{\mathbf{b}}$  by,

$$(\mathbf{A})_{\text{index}(i, j), \text{index}(k, j)} = {}_L p_{B, N_y, i, k}^{(2)}, \quad k = 0, \dots, N_y, k \neq i, \quad (4.19a)$$

$$(\mathbf{A})_{\text{index}(i, j), \text{index}(i, k)} = -t_f p_{B, N_t, j, k}^{(1)}, \quad k = 0, \dots, N_t, k \neq j, \quad (4.19b)$$

$$(\mathbf{A})_{\text{index}(i, j), \text{index}(N_y+1, k)} = t_f p_{B, N_t, j, k}^{(1)}, \quad k = 0, \dots, N_t, \quad (4.19c)$$

$$(\mathbf{A})_{\text{index}(i, j), \text{index}(i, j)} = {}_L p_{B, N_y, i, i}^{(2)} - t_f p_{B, N_t, j, j}^{(1)}, \quad (4.19d)$$

$$(\mathbf{B})_{\text{index}(i, j), \text{index}(i, k)} = -t_f p_{B, N_t, j, k}^{(1)}, \quad k = 0, \dots, N_t, \quad (4.19e)$$

$$(\mathbf{B})_{\text{index}(i, j), \text{index}(N_y+1, k)} = t_f p_{B, N_t, j, k}^{(1)}, \quad k = 0, \dots, N_t, \quad (4.19f)$$

$$\left( \tilde{\mathbf{b}} \right)_n = (\mathbf{F})_i, \quad (4.19g)$$

for  $i = 0, \dots, N_y$ ;  $j = 0, \dots, N_t$ . Therefore, we can write Eqs. (4.9) in matrix form as,

$$\bar{\mathbf{A}} \hat{\phi} + \bar{\mathbf{B}} \hat{u} = \bar{\mathbf{b}}, \quad (4.20)$$

where,

$$\left( (\mathbf{A})_{i, *} : (i+1) \not\equiv 0 \pmod{(N_y+2)} \wedge i \in \{0, \dots, L''\} \right) = \bar{\mathbf{A}} \in \mathbb{R}^{(L'+1) \times (L''+2)}, \quad (4.21)$$

$$\left( (\mathbf{B})_{i, *} : (i+1) \not\equiv 0 \pmod{(N_y+2)} \wedge i \in \{0, \dots, L''\} \right) = \bar{\mathbf{B}} \in \mathbb{R}^{(L'+1) \times (L''+2)}, \quad (4.22)$$

$$\left( \left( \tilde{\mathbf{b}} \right)_i : (i+1) \not\equiv 0 \pmod{(N_y+2)} \wedge i \in \{0, \dots, L''\} \right) = \bar{\mathbf{b}} \in \mathbb{R}^{(L'+1)}, \quad (4.23)$$

and the asterisk “\*” denotes the whole range of column indices. Hence, the global collocation matrix  $\bar{\mathbf{H}}$  is simply given by,

$$\bar{\mathbf{H}} = [\bar{\mathbf{A}}, \bar{\mathbf{B}}], \quad (4.24)$$

where “[.,.]” is the usual horizontal matrix concatenation notation. Thus, Eq. (4.20) can be rewritten as,

$$\bar{\mathbf{H}} \mathbf{Z} = \bar{\mathbf{b}}, \quad (4.25)$$

where the solution vector  $\mathbf{Z} \in \mathbb{R}^{2L''+4}$  is given by,

$$\mathbf{Z} = \text{vec} \begin{bmatrix} \hat{\phi}, \hat{\mathbf{u}} \end{bmatrix}, \quad (4.26)$$

and “vec” denotes the vectorization of a matrix. Moreover, if we define the elements of the column vector  $\bar{\phi} \in \mathbb{R}^{L'+1}$  by,

$$\bar{\phi}_l = \hat{\phi}_{l+\lfloor l/(N_y+1) \rfloor}, \quad l = 0, \dots, L', \quad (4.27)$$

then Eqs. (4.12) can also be written in the following useful matrix form,

$$\left( \mathbf{I}_{N_t+1} \otimes {}^L \mathbf{P}_{B, N_y, N_y+2}^{(1)} \right) \bar{\phi} = \mathbf{0} \in \mathbb{R}^{N_t+1}, \quad (4.28)$$

where  $\mathbf{I}_{N_t+1}$  is the identity matrix of order  $N_t + 1$ . We can further combine Eqs. (4.25) and (4.28) in a single linear system form. To this end, define the index vector  $\Lambda$  by,

$$\Lambda = [0(1)L'] + \left\lfloor \frac{1}{N_y+1} [0(1)L'] \right\rfloor, \quad (4.29)$$

where  $[0(1)L'] = [0, 1, \dots, L']$ , and “ $\lfloor \cdot \rfloor$ ” denotes the floor function. Moreover, let

$$\Psi \in \mathbb{R}^{(N_t+1) \times (2L''+4)} : \begin{cases} (\Psi)_{i,\Lambda} = \left( I_{N_t+1} \otimes {}^L P_{B, N_y, N_y+2}^{(1)} \right)_{i,*}, & i = 0, \dots, N_t, \\ (\Psi)_{i,j} = 0, & i = 0, \dots, N_t; j \notin \Lambda. \end{cases} \quad (4.30)$$

Hence, Eqs. (4.25) and (4.28) can be written as,

$$\mathbf{H} \mathbf{Z} = \mathbf{b}, \quad (4.31)$$

where,

$$\mathbf{H} = [\bar{\mathbf{H}}; \Psi], \quad (4.32)$$

$$\mathbf{b} = \text{vec} [\bar{\mathbf{b}}, \mathbf{0}], \quad (4.33)$$

$\mathbf{0} \in \mathbb{R}^{N_t+1}$ , and “ $[\cdot; \cdot]$ ” is the vertical matrix concatenation along columns defined by “ $[\cdot^T, \cdot^T]^T$ ”. To write the discrete cost functional in terms of the solution vector  $\mathbf{Z}$ , we introduce the mapping  $m = \text{index}(k, l) : m = k + l(N_y + 2)$ , and the notation,

$$(\mathbf{v})_{(k)} = \underbrace{\mathbf{v} \circ \mathbf{v} \circ \dots \circ \mathbf{v}}_{k\text{-times}} \forall \mathbf{v} \in \mathbb{R}^l, l \in \mathbb{Z}^+, \quad (4.34)$$

where “ $\circ$ ” denotes the Hadamard (entrywise) product. Moreover, let  $\check{\phi} = (\check{\phi}_i) \in \mathbb{R}^{L'+1}$  and  $\check{\mathbf{u}} = (\check{u}_i) \in \mathbb{R}^{L'+1}$ :

$$\check{\phi}_i = (\mathbf{Z})_{\lfloor \frac{i}{N_t+1} \rfloor + (N_y+2) \cdot \left( 1 - \delta_{\lfloor \frac{i}{N_t+1} \rfloor, \lfloor \frac{i}{N_t+1} \rfloor} \right) \cdot i \equiv 0 \pmod{(N_t+1)}}, \quad i = 0, \dots, L', \quad (4.35)$$

$$\check{u}_i = (\mathbf{Z})_{L'' + \lfloor \frac{i}{N_t+1} \rfloor + (N_y+2) \cdot \left( 1 - \delta_{\lfloor \frac{i}{N_t+1} \rfloor, \lfloor \frac{i}{N_t+1} \rfloor} \right) \cdot i \equiv 0 \pmod{(N_t+1)} + 2}, \quad i = 0, \dots, L', \quad (4.36)$$

then the sought discrete cost functional can be written as,

$$J_{N_y, N_t=t_f} \mathbf{P}_{B, N_t, N_t+2}^{(1)} \left( I_{N_t+1} \otimes {}^L \mathbf{P}_{B, N_y, N_y+2}^{(1)} \right) \left( r_1 (\bar{\mathbf{x}})_{(2)} + r_2 (\bar{\mathbf{u}})_{(2)} \right), \quad (4.37)$$

where,

$$(\bar{\mathbf{u}})_q = \hat{u}_{q + \lfloor \frac{q}{N_y+1} \rfloor} = (\mathbf{Z})_{L'' + q + \lfloor \frac{q}{N_y+1} \rfloor + 2}, \quad q = 0, \dots, L', \quad (4.38)$$

$$(\bar{\mathbf{x}})_q = \hat{x}_{q + \lfloor \frac{q}{N_y+1} \rfloor}, \quad q = 0, \dots, L', \quad (4.39)$$

$$\hat{x}_m = x_{k,l} = \left[ I_{N_y+1} \otimes {}^L \mathbf{P}_{B, N_t, l}^{(1)} \right]_{k,*} (\check{\phi} + \check{\mathbf{u}}) + f_k, \quad m = 0, \dots, L'' + 1. \quad (4.40)$$

Hence, the optimal control problem has been reduced to a nonlinear programming problem, in which we seek the minimization of the objective function  $J_{N_y, N_t}$  defined by (4.37) subject to the linear constraints given by (4.31). Solving for  $\mathbf{Z}$  yields the values of the functions  $\phi$  and  $u$  at the solution nodes  $\left(y_{L, N_y, i}^{(\alpha)}, t_{t_f, N_t, j}^{(\alpha)}\right), i = 0, \dots, N_y + 1; j = 0, \dots, N_t$ . To recover the state function  $x$  at those nodes, we can use Eq. (3.7) to obtain,

$$x_{i,j} \approx {}_{t_f} \mathbf{P}_{B, N_t, j}^{(1)} \left( \phi \left( y_{L, N_y, i}^{(\alpha)}, t_{t_f, N_t}^{(\alpha)} \right) + u \left( y_{L, N_y, i}^{(\alpha)}, t_{t_f, N_t}^{(\alpha)} \right) \right) + f_i \forall i, j. \quad (4.41)$$

Furthermore, we can generate the approximation of the state profile on  $D_{L, t_f}^2$  using the polynomial interpolant of  $x$  given by,

$$P_{N_y, N_t} x(y, t) = \sum_{s=0}^{N_y} \sum_{k=0}^{N_t} x_{s,k} {}_{L, t_f} \mathcal{L}_{N_y, N_t, s, k}^{(\alpha)}(y, t). \quad (4.42)$$

## 5. ERROR ANALYSIS OF THE BSGPM

Let

$$I_{q, \tilde{x}}^{(x)}(f(x)) = \int_0^{\tilde{x}} \int_0^{\sigma_{q-1}} \dots \int_0^{\sigma_2} \int_0^{\sigma_1} f(\sigma_0) d\sigma_0 d\sigma_1 \dots d\sigma_{q-2} d\sigma_{q-1}, \quad (5.1)$$

be the  $q$ -fold integral of any integrable single-variable function  $f(x)$ , for some positive real number  $\tilde{x} \in [0, l], l \in \mathbb{R}^+$ , and denote  $\mathbb{Z}^+ \cup \{0\}$  by  $\mathbb{Z}_0^+$ . The following two theorems highlight the truncation error and the error bounds of the barycentric shifted Gegenbauer quadrature (BSGQ) associated with the BSGIM,  ${}_l \mathbf{P}_{B, n}^{(1)} \forall n \in \mathbb{Z}^+$ . The proof of both theorems follow that of [Elgindy (2016a), Theorems 4.1 & 4.3].

*Theorem 5.1*

Let  $f(x) \in C^{n+1}[0, l]$ , be interpolated by the shifted Gegenbauer polynomials at the SGG nodes,  $x_{l, n, i}^{(\alpha)} \in \mathbb{S}_{l, n}^{(\alpha)}, i = 0, \dots, n$ , then there exist some numbers  $\zeta_{l, n, i}^{(\alpha)} = \zeta \left( x_{l, n, i}^{(\alpha)} \right) \in (0, l), i = 0, \dots, n$ , such that,

$$I_{1, x_{l, n, i}^{(\alpha)}}^{(x)}(f(x)) = {}_l \mathbf{P}_{B, n, i}^{(1)} \mathbf{f} + E_{1, l, n}^{(\alpha)} \left( x_{l, n, i}^{(\alpha)}, \zeta_{l, n, i}^{(\alpha)} \right), \quad (5.2)$$

where  $\mathbf{f} = f \left( \mathbf{x}_{l, n}^{(\alpha)} \right), \mathbf{x}_{l, n}^{(\alpha)} = \left[ x_{l, n, 0}^{(\alpha)}, x_{l, n, 1}^{(\alpha)}, \dots, x_{l, n, n}^{(\alpha)} \right]^T$ ,

$$E_{1, l, n}^{(\alpha)} \left( x_{l, n, i}^{(\alpha)}, \zeta_{l, n, i}^{(\alpha)} \right) = \frac{f^{(n+1)} \left( \zeta_{l, n, i}^{(\alpha)} \right)}{(n+1)! K_{l, n+1}^{(\alpha)}} I_{1, x_{l, n, i}^{(\alpha)}}^{(x)} \left( G_{l, n+1}^{(\alpha)}(x) \right), \quad (5.3)$$

and  $K_{l, n}^{(\alpha)}$  is the leading coefficient of the shifted Gegenbauer polynomial  $G_{l, n}^{(\alpha)}(x)$ .

*Theorem 5.2*

Assume that  $f(x) \in C^{n+1}[0, l]$ , and

$$\left\| f^{(n+1)} \right\|_{L^\infty[0, l]} = \sup_{0 \leq x \leq l} |f(x)| \leq A \in \mathbb{R}^+,$$

for some number  $n \in \mathbb{Z}_0^+$ . Moreover, let  $I_{1, x_{l, n, i}^{(\alpha)}}^{(x)}(f(x))$ , be approximated by the BSGQ, for each integration node  $x_{l, n, i}^{(\alpha)} \in \mathbb{S}_{l, n}^{(\alpha)}, i = 0, \dots, n$ . Then there exist some positive constants  $D_1^{(\alpha)}$  and  $D_2^{(\alpha)}$ , independent of  $n$  such that the truncation error of the BSGQ,  $E_{1, l, n}^{(\alpha)}$ , is bounded by the following inequalities:

$$\left| E_{1, l, n}^{(\alpha)} \right| \leq \frac{A 2^{-2n-1} \Gamma(\alpha+1) x_i l^{n+1} \Gamma(n+2\alpha+1)}{\Gamma(2\alpha+1) \Gamma(n+2) \Gamma(n+\alpha+1)} \left( \left\{ \begin{array}{l} 1, \quad n \geq 0 \wedge \alpha \geq 0, \\ \frac{(n+1)! \Gamma(2\alpha)}{\Gamma(n+2\alpha+1)} \left| \left( \frac{n+1}{2} + \alpha - 1 \right) \right|, \quad \frac{n+1}{2} \in \mathbb{Z}^+ \wedge -\frac{1}{2} < \alpha < 0 \end{array} \right\} \right), \quad (5.4)$$

$$\left| E_{1,l,n}^{(\alpha)} \right| < \frac{A 2^{-2n-1} \Gamma(\alpha) |\alpha| x_i l^{n+1}}{\sqrt{(n+1)(2\alpha+n+1)} \Gamma(n+\alpha+1)} \left| \left( \frac{n}{2} + \alpha \right) \right|, \quad \frac{n}{2} \in \mathbb{Z}_0^+ \wedge -\frac{1}{2} < \alpha < 0, \quad (5.5)$$

$$\left| E_{1,l,n}^{(\alpha)} \right| \leq \begin{cases} B_1^{(\alpha)} \frac{e^n l^{n+1} x_i}{2^{2n+1} n^{n+3/2-\alpha}}, & \alpha \geq 0 \wedge n \gg 1, \\ B_2^{(\alpha)} \frac{e^n l^{n+1} x_i}{2^{2n+1} n^{n+3/2}}, & -\frac{1}{2} < \alpha < 0 \wedge n \gg 1, \end{cases} \quad (5.6a)$$

$$(5.6b)$$

for all  $i = 0, \dots, n$ , where  $B_1^{(\alpha)} = A D_1^{(\alpha)}$ , and  $B_2^{(\alpha)} = B_1^{(\alpha)} D_2^{(\alpha)}$ .

Theorems 5.1 and 5.2 show that the truncation error associated with the BSGQ decays exponentially fast for increasing values of  $n$  with  $\left| E_{1,l,n}^{(0)} \right| < \left| E_{1,l,n}^{(\alpha)} \right| \forall \alpha \neq 0$ , as  $n \rightarrow \infty$ .

The following two theorems further generalize Theorems 5.1 and 5.2 for higher-order BSGQs.

**Theorem 5.3**

Let  $f(x) \in C^{n+1}[0, l]$ , be interpolated by the shifted Gegenbauer polynomials at the SGG nodes,  $x_{l,n,i}^{(\alpha)} \in \mathbb{S}_{l,n}^{(\alpha)}$ ,  $i = 0, \dots, n$ , then there exist some numbers  $\zeta_{l,n,i}^{(\alpha)} = \zeta \left( x_{l,n,i}^{(\alpha)} \right) \in (0, l)$ ,  $i = 0, \dots, n$ , such that,

$$I_{q,x_{l,n,i}^{(\alpha)}}^{(x)}(f(x)) = {}_l \mathbf{P}_{B,n,i}^{(q)} \mathbf{f} + E_{q,l,n}^{(\alpha)} \left( x_{l,n,i}^{(\alpha)}, \zeta_{l,n,i}^{(\alpha)} \right), \quad (5.7)$$

$$E_{q,l,n}^{(\alpha)} \left( x_{l,n,i}^{(\alpha)}, \zeta_{l,n,i}^{(\alpha)} \right) = \frac{1}{(q-1)! (n+1)! K_{l,n+1}^{(\alpha)}} I_{1,x_{l,n,i}^{(\alpha)}}^{(x)} \left( G_{l,n+1}^{(\alpha)}(x) \right) \left[ \frac{d^{n+1}}{dx^{n+1}} \left( \left( x_{l,n,i}^{(\alpha)} - x \right)^{q-1} f(x) \right) \right]_{x=\zeta_{l,n,i}^{(\alpha)}}. \quad (5.8)$$

**Proof**

Using the above notation, we can write Cauchy's formula for repeated integration in the following simple form,

$$I_{q,x_{l,n,i}^{(\alpha)}}^{(x)}(f(x)) = \frac{1}{(q-1)!} I_{1,x_{l,n,i}^{(\alpha)}}^{(x)} \left( \left( x_{l,n,i}^{(\alpha)} - x \right)^{q-1} f(x) \right). \quad (5.9)$$

By Theorem 5.1,

$$I_{1,x_{l,n,i}^{(\alpha)}}^{(x)} \left( \left( x_{l,n,i}^{(\alpha)} - x \right)^{q-1} f(x) \right) = {}_l \mathbf{P}_{B,n,i}^{(1)} \bar{\mathbf{f}} + \frac{1}{(n+1)! K_{l,n+1}^{(\alpha)}} I_{1,x_{l,n,i}^{(\alpha)}}^{(x)} \left( G_{l,n+1}^{(\alpha)}(x) \right) \left[ \frac{d^{n+1}}{dx^{n+1}} \left( \left( x_{l,n,i}^{(\alpha)} - x \right)^{q-1} f(x) \right) \right]_{x=\zeta_{l,n,i}^{(\alpha)}}, \quad (5.10)$$

where,

$$\bar{\mathbf{f}} = \left( x_{l,n,i}^{(\alpha)} \mathbf{1} - x_{l,n}^{(\alpha)} \right)_{(q-1)} \circ \mathbf{f}, \quad (5.11)$$

and  $\mathbf{1} \in \mathbb{R}^{n+1}$ , is the all-ones vector. Hence, Theorem 5.3 follows directly by noticing that,

$$\frac{1}{(q-1)!} {}_l \mathbf{P}_{B,n,i}^{(1)} \left( \left( x_{l,n,i}^{(\alpha)} \mathbf{1} - x_{l,n}^{(\alpha)} \right)_{(q-1)} \circ \mathbf{f} \right) = \frac{1}{(q-1)!} \left( \left( x_{l,n,i}^{(\alpha)} \mathbf{1} - x_{l,n}^{(\alpha)} \right)_{(q-1)}^T \circ {}_l \mathbf{P}_{B,n,i}^{(1)} \right) \mathbf{f} = {}_l \mathbf{P}_{B,n,i}^{(q)} \mathbf{f}. \quad (5.12)$$

□

**Theorem 5.4**

Assume that  $f(x) \in C^{n+1}[0, l]$ , and

$$\sup_{0 < x < l} \left| \frac{d^{n+1}}{dx^{n+1}} \left( \left( x_{l,n,i}^{(\alpha)} - x \right)^{q-1} f(x) \right) \right| \leq A \in \mathbb{R}^+, \quad (5.13)$$

for some numbers  $n \in \mathbb{Z}_0^+$ , and  $q \in \mathbb{Z}^+$ . Moreover, let  $I_{q,x_{l,n,i}^{(\alpha)}}^{(x)}(f(x))$ , be approximated by the  $q$ th-order BSGQ, for each integration node  $x_{l,n,i}^{(\alpha)} \in \mathbb{S}_{l,n}^{(\alpha)}$ ,  $i = 0, \dots, n$ . Then there exist some positive constants  $D_1^{(\alpha)}$  and  $D_2^{(\alpha)}$ , independent of  $n$  such that the truncation error of the  $q$ th-order BSGQ,  $E_{q,l,n}^{(\alpha)}$ , is bounded by the following



inequalities:

$$\left| E_{q,l,n}^{(\alpha)} \right| \leq \frac{A 2^{-2n-1} \Gamma(\alpha+1) x_i l^{n+1} \Gamma(n+2\alpha+1)}{(q-1)! \Gamma(2\alpha+1) \Gamma(n+2) \Gamma(n+\alpha+1)} \left( \left\{ \begin{array}{l} 1, \quad n \geq 0 \wedge \alpha \geq 0, \\ \frac{(n+1)! \Gamma(2\alpha)}{\Gamma(n+2\alpha+1)} \left| \left( \frac{\frac{n+1}{2} + \alpha - 1}{\frac{n+1}{2}} \right) \right|, \quad \frac{n+1}{2} \in \mathbb{Z}^+ \wedge -\frac{1}{2} < \alpha < 0 \end{array} \right. \right), \quad (5.14)$$

$$\left| E_{q,l,n}^{(\alpha)} \right| < \frac{A 2^{-2n-1} \Gamma(\alpha) |\alpha| x_i l^{n+1}}{(q-1)! \sqrt{(n+1)(2\alpha+n+1)} \Gamma(n+\alpha+1)} \left| \left( \frac{\frac{n}{2} + \alpha}{\frac{n}{2}} \right) \right|, \quad \frac{n}{2} \in \mathbb{Z}_0^+ \wedge -\frac{1}{2} < \alpha < 0, \quad (5.15)$$

$$\left| E_{q,l,n}^{(\alpha)} \right| \leq \begin{cases} B_1^{(\alpha)} \frac{e^n l^{n+1} x_i}{2^{2n+1} n^{n+3/2-\alpha} (q-1)!}, & \alpha \geq 0 \wedge n \gg 1, \\ B_2^{(\alpha)} \frac{e^n l^{n+1} x_i}{2^{2n+1} n^{n+3/2} (q-1)!}, & -\frac{1}{2} < \alpha < 0 \wedge n \gg 1, \end{cases} \quad (5.16a)$$

$$(5.16b)$$

for all  $i = 0, \dots, n$ , where  $B_1^{(\alpha)} = A D_1^{(\alpha)}$ , and  $B_2^{(\alpha)} = B_1^{(\alpha)} D_2^{(\alpha)}$ .

*Proof*

The proof follows easily from Theorems 5.1 and 5.3.  $\square$

The following is a direct corollary of Theorem 5.3.

*Corollary 5.1*

Let  $f(x) \in C^{n+1}[0, l]$ , be interpolated by the shifted Gegenbauer polynomials at the SGG nodes,  $x_{l,n,i}^{(\alpha)} \in \mathbb{S}_{l,n}^{(\alpha)}$ ,  $i = 0, \dots, n$ , and suppose that,

$$\left\| f^{(k)} \right\|_{L^\infty[0,l]} \leq A_{\max} \in \mathbb{R}^+ \forall k = 0, \dots, n+1. \quad (5.17)$$

Then there exist some numbers  $\zeta_{l,n,i}^{(\alpha)} = \zeta \left( x_{l,n,i}^{(\alpha)} \right) \in (0, l)$ ,  $i = 0, \dots, n$ , such that,

$$I_{q,x_{l,n,i}^{(\alpha)}}^{(x)}(f(x)) = {}_l \mathbf{P}_{B,n,i}^{(q)} \mathbf{f} + E_{q,l,n}^{(\alpha)} \left( x_{l,n,i}^{(\alpha)}, \zeta_{l,n,i}^{(\alpha)} \right), \quad (5.18)$$

$$\left| E_{q,l,n}^{(\alpha)} \left( x_{l,n,i}^{(\alpha)}, \zeta_{l,n,i}^{(\alpha)} \right) \right| \leq \frac{2^{-n} A_{\max} l^{1+n} n_{\max,i} x_{l,n,i}^{(\alpha)} \Gamma(1+\alpha) \Gamma(1+n+2\alpha)}{(n+1)! (q-1)! \Gamma(1+n+\alpha) \Gamma(1+2\alpha)} \times \left( \left\{ \begin{array}{l} 1, \quad \alpha \geq 0, \\ 2|\alpha| \left| \left( \frac{\frac{n}{2} + \alpha}{\frac{n}{2}} \right) \right|^{(1+n)! \Gamma(2\alpha)} \\ \frac{\sqrt{(1+n)(1+n+2\alpha)} \Gamma(1+n+2\alpha)}{\left| \left( \frac{\frac{n-1}{2} + \alpha}{\frac{n+1}{2}} \right) \right|^{(n+1)! \Gamma(2\alpha)}}, \quad \frac{n}{2} \in \mathbb{Z}_0^+ \wedge -\frac{1}{2} < \alpha < 0, \\ \frac{\left| \left( \frac{\frac{n-1}{2} + \alpha}{\frac{n+1}{2}} \right) \right|^{(n+1)! \Gamma(2\alpha)}}{\Gamma(1+n+2\alpha)}, \quad \frac{n+1}{2} \in \mathbb{Z}^+ \wedge -\frac{1}{2} < \alpha < 0 \end{array} \right. \right), \quad (5.19)$$

for some positive number  $q \in \mathbb{Z}^+$ , where,

$$n_{\max,i} = \max \left\{ (q-1)!, \sup_{0 \leq k \leq n+1} \left( (q-1)_{n-k+1} \left| x_{l,n,i}^{(\alpha)} - \zeta_{l,n,i}^{(\alpha)} \right|^{q-n+k-2} \right) \right\}, \quad (5.20)$$

and,

$$(x)_n = x(x-1)\dots(x-(n-1)),$$

is the falling factorial.

*Proof*

To simplify the notation, let

$$\mu(x) = \frac{d^{n+1}}{dx^{n+1}} \left( \left( x_{l,n,i}^{(\alpha)} - x \right)^{q-1} f(x) \right).$$

Then by the general Leibniz rule, we have

$$(\mu(x) f(x))^{(n+1)} = \sum_{k=0}^{n+1} \binom{n+1}{k} \mu^{(n-k+1)}(x) f^{(k)}(x).$$

Realizing that,

$$\left| \mu^{(n-k+1)}(\zeta_{l,n,i}^{(\alpha)}) \right| \leq \begin{cases} 0, & n-k > q-2, \\ (q-1)!, & n-k = q-2, \\ (q-1)_{n-k+1} \left| x_{l,n,i}^{(\alpha)} - \zeta_{l,n,i}^{(\alpha)} \right|^{q-n+k-2}, & n-k < q-2. \end{cases},$$

we find that,

$$\left| E_{q,l,n}^{(\alpha)} \left( x_{l,n,i}^{(\alpha)}, \zeta_{l,n,i}^{(\alpha)} \right) \right| \leq \frac{2^{n+1} n_{\max,i} A_{\max}}{(q-1)! (n+1)! \left| K_{l,n+1}^{(\alpha)} \right|} \left| I_{1,x_{l,n,i}^{(\alpha)}}^{(x)} \left( G_{l,n+1}^{(\alpha)}(x) \right) \right|. \quad (5.21)$$

The corollary follows easily using [Elgindy (2016a), Lemma 4.1].  $\square$

Using the above error analysis, we can straightforwardly determine the truncation error of the integral dynamical system equation (3.10) as stated in the following theorem.

*Theorem 5.5*

Let

$$\psi(y, t) = \phi(0, t) - \phi(y, t) + u(0, t) - u(y, t) \forall (y, t) \in D_{L,t_f}^2. \quad (5.22)$$

Suppose also that  $\psi(y, t) \in C^{N_t+1}[0, t_f] \forall y \in [0, L]$ , and,

$$\left\| \frac{\partial^k \phi}{\partial y^k} \right\|_{L^\infty(D_{L,t_f}^2)} \leq A_{\max} \in \mathbb{R}^+ \forall k = 0, \dots, N_y + 1. \quad (5.23)$$

Then there exist some numbers  $\zeta_{L,N_y,i}^{(\alpha),y} = \zeta \left( y_{L,N_y,i}^{(\alpha)} \right) \in (0, L)$ ,  $\zeta_{t_f,N_t,j}^{(\alpha),t} = \zeta \left( t_{t_f,N_t,j}^{(\alpha)} \right) \in (0, t_f)$ ,  $i = 0, \dots, N_y$ ;  $j = 0, \dots, N_t$ , such that the BSGPM discretizes the integral dynamical system equation (3.10) with a total truncation error,  $E_{\text{total},i,j}$  at each SGG point,  $\left( y_{L,N_y,i}^{(\alpha)}, t_{t_f,N_t,j}^{(\alpha)} \right)$ , bounded by,

$$|E_{\text{total},i,j}| \leq D^{(\alpha)} (\varepsilon_{1,j} + \varepsilon_{2,i}), \quad (5.24)$$

where,

$$\varepsilon_{1,j} = \frac{4^{-N_t} B_{\max,j} t_f^{N_t+1} t_{t_f,N_t,j}^{(\alpha)}}{(N_t+1)! \Gamma(N_t+\alpha+1)} \left( \begin{cases} \Gamma(N_t+2\alpha+1), & \alpha \geq 0, \\ 2\alpha \left( \frac{\frac{N_t}{2} + \alpha}{\frac{N_t}{2}} \right)_{(N_t+1)! \Gamma(2\alpha)} \\ - \frac{\sqrt{(N_t+1)(N_t+2\alpha+1)}}{\left| \left( \frac{\frac{N_t-1}{2} + \alpha}{\frac{N_t+1}{2}} \right) \right| (N_t+1)! \Gamma(2\alpha)}, & \frac{N_t}{2} \in \mathbb{Z}_0^+ \wedge -\frac{1}{2} < \alpha < 0, \\ \left| \left( \frac{\frac{N_t-1}{2} + \alpha}{\frac{N_t+1}{2}} \right) \right| (N_t+1)! \Gamma(2\alpha), & \frac{N_t+1}{2} \in \mathbb{Z}^+ \wedge -\frac{1}{2} < \alpha < 0 \end{cases} \right), \quad (5.25)$$

$$\varepsilon_{2,i} = \frac{A_{\max} n_{\max,i} 2^{1-N_y} y_{L,N_y,i}^{(\alpha)} L^{N_y+1} \Gamma(N_y+2\alpha+1)}{(N_y+1)! \Gamma(1+N_y+\alpha)} \left( \begin{cases} 1, & \alpha \geq 0, \\ 2|\alpha| \left| \left( \frac{\frac{N_y}{2} + \alpha}{\frac{N_y}{2}} \right) \right|_{(N_y+1)! \Gamma(2\alpha)} \\ \frac{\sqrt{(N_y+1)(N_y+2\alpha+1)} \Gamma(N_y+2\alpha+1)}{\left| \left( \frac{\frac{N_y-1}{2} + \alpha}{\frac{N_y+1}{2}} \right) \right| (N_y+1)! \Gamma(2\alpha)}, & \frac{N_y}{2} \in \mathbb{Z}_0^+ \wedge -\frac{1}{2} < \alpha < 0, \\ \frac{\left| \left( \frac{\frac{N_y-1}{2} + \alpha}{\frac{N_y+1}{2}} \right) \right| (N_y+1)! \Gamma(2\alpha)}{\Gamma(N_y+2\alpha+1)}, & \frac{N_y+1}{2} \in \mathbb{Z}^+ \wedge -\frac{1}{2} < \alpha < 0 \end{cases} \right), \quad (5.26)$$

$$n_{\max,i} = \max \left\{ 1, \sup \left| y_{L,N_y,i}^{(\alpha)} - \zeta_{L,N_y,i}^{(\alpha),y} \right| \right\}, \quad (5.27)$$

$$D^{(\alpha)} = \frac{\Gamma(1+\alpha)}{2\Gamma(1+2\alpha)}, \quad (5.28)$$

assuming that,

$$\left\| \left[ \frac{\partial^{N_t+1}}{\partial t^{N_t+1}} \psi(y, t) \right]_{t=\zeta_{t_f, N_t, j}^{(\alpha), t}} \right\|_{L^\infty[0, L]} \leq B_{\max, j} \in \mathbb{R}^+ \forall j. \quad (5.29)$$

*Proof*

A straightforward error analysis shows that,

$$|E_{\text{total}, i, j}| \leq \frac{1}{(N_t + 1)! |K_{t_f, N_t+1}^{(\alpha)}|} \left| I_{1, t_{t_f, N_t, j}}^{(t)} \left( G_{t_f, N_t+1}^{(\alpha)}(t) \right) \left[ \frac{\partial^{N_t+1}}{\partial t^{N_t+1}} \psi(y, t) \right]_{t=\zeta_{t_f, N_t, j}^{(\alpha), t}} \right| \quad (5.30)$$

$$+ \frac{2^{N_y+1} n_{\max, i} A_{\max}}{(N_y + 1)! |K_{L, N_y+1}^{(\alpha)}|} \left| I_{1, y_{L, N_y, i}}^{(y)} \left( G_{L, N_y+1}^{(\alpha)}(y) \right) \right|, \quad (5.31)$$

from which the theorem follows.  $\square$

## 6. NUMERICAL EXAMPLE

In this section, we report the results of the developed BSGPM on the optimal control problem under study with  $L = 4, t_f = 1, r_1 = r_2 = 1/2; f(y) = 1 + y$ . The BSGPM was applied using  $N_y = N_t = N = 4, 5, \dots, 12$ , and  $\alpha = -0.4, -0.3, \dots, 0.9$ . The nonlinear programming problem was solved using MATLAB “fmincon” constrained optimization solver with the default “TolFun” and “TolCon” of  $10^{-6}$ . The numerical experiments were conducted on a personal laptop equipped with an Intel(R) Core(TM) i7-2670QM CPU with 2.20GHz speed running on a Windows 10 64-bit operating system and provided with MATLAB R2014b (8.4.0.150421) Software. Figure 1 shows the plots of the approximate optimal cost functional  $J_{N,N}$ , the feasibility<sup>†</sup> of the optimal solution  $Z^*$  as reported by the solver, the maximum error in the initial condition (2.3),  $\psi_1$ , at the 101 linearly spaced nodes in the  $y$ - and  $t$ -directions from 0 to 4, and 0 to 1, respectively depicted in semi-logarithmic scale, and the maximum error in the boundary condition (3.6),  $\psi_2$ . As can be observed from the figure, the approximate optimal cost functional  $J_{N,N}^*$  is approximately 15 for all input data with feasibility and  $\psi_2$  near the machine epsilon. We observe also that discretizations at the SGG points for non-positive  $\alpha$ -values yield the minimum  $\psi_1$ -values for small values of  $N$ , whereas the accuracy degrades for increasing values of  $\alpha$ —a result that is consistent with the work of [Elgindy (2016a)] on second-order one-dimensional hyperbolic telegraph equations. We expect also to obtain the optimal approximations in the maximum norm for large values of  $N$  through discretizations at the shifted Chebyshev-Gauss points as discussed earlier in Section 5. Fortunately, the present numerical scheme converges exponentially fast for sufficiently smooth solutions using relatively small number of grids. A sketch of the calculated state and control profiles using  $N = 12$  and  $\alpha = -0.2$  is shown in Figure 2. Figure 3 shows also their profiles at the midpoint  $y = 2$ . In comparison with [Rad et al. (2014)] who solved the optimal control problem using 120 nodal points in both directions, the BSGPM exhibits exponential convergence rates and produces excellent approximations using as small as 5 nodes in both directions.

## 7. CONCLUSION

This paper presented a robust and computationally efficient BSGPM for solving a PDE-governed optimal control problem. A key reason underlying the computationally streamlined nature of the current approach lies in the accurate discretization of the system dynamics and constraints into a well-conditioned algebraic linear system using

<sup>†</sup>By the feasibility of the solution, we mean the maximum constraint violation.

stable and high-order BSGQs. Using a practical test example, it is shown that the BSGPM has two significant advantages over the method of [Rad et al. (2014)]: (i) the method converges exponentially fast, and (ii) the required number of collocation/nodal points to produce high-order approximations is significantly smaller. The test example also suggests that discretizations at the SGG points for non-positive  $\alpha$ -values yield better approximations for relatively small numbers of expansion terms, whereas the accuracy degrades for increasing values of  $\alpha$ . The present method provides a strong addition to the arsenal of numerical pseudospectral methods, and can be extended to solve a wide range of PDE-governed optimal control problems arising in numerous applications.

## REFERENCES

- [Elgindy (2013)]. Elgindy, K., 2013. Gegenbauer Collocation Integration Methods: Advances in Computational Optimal Control Theory. Ph.D. thesis, School of Mathematical Sciences, Faculty of Science, Monash University.
- [Elgindy (2016a)]. Elgindy, K. T., January 2016a. High-order numerical solution of second-order one-dimensional hyperbolic telegraph equation using a shifted Gegenbauer pseudospectral method. *Numerical Methods for Partial Differential Equations* 32 (1), 307–349.
- [Elgindy (2016b)]. Elgindy, K. T., 2016b. Optimal barycentric Gegenbauer quadrature, arXiv:1603.00945 [math.NA].
- [Rad et al. (2014)]. Rad, J., Kazem, S., Parand, K., 2014. Optimal control of a parabolic distributed parameter system via radial basis functions. *Communications in Nonlinear Science and Numerical Simulation* 19 (8), 2559–2567.
- [Trefethen (2000)]. Trefethen, L. N., 2000. *Spectral Methods in MATLAB*. SIAM, Philadelphia.

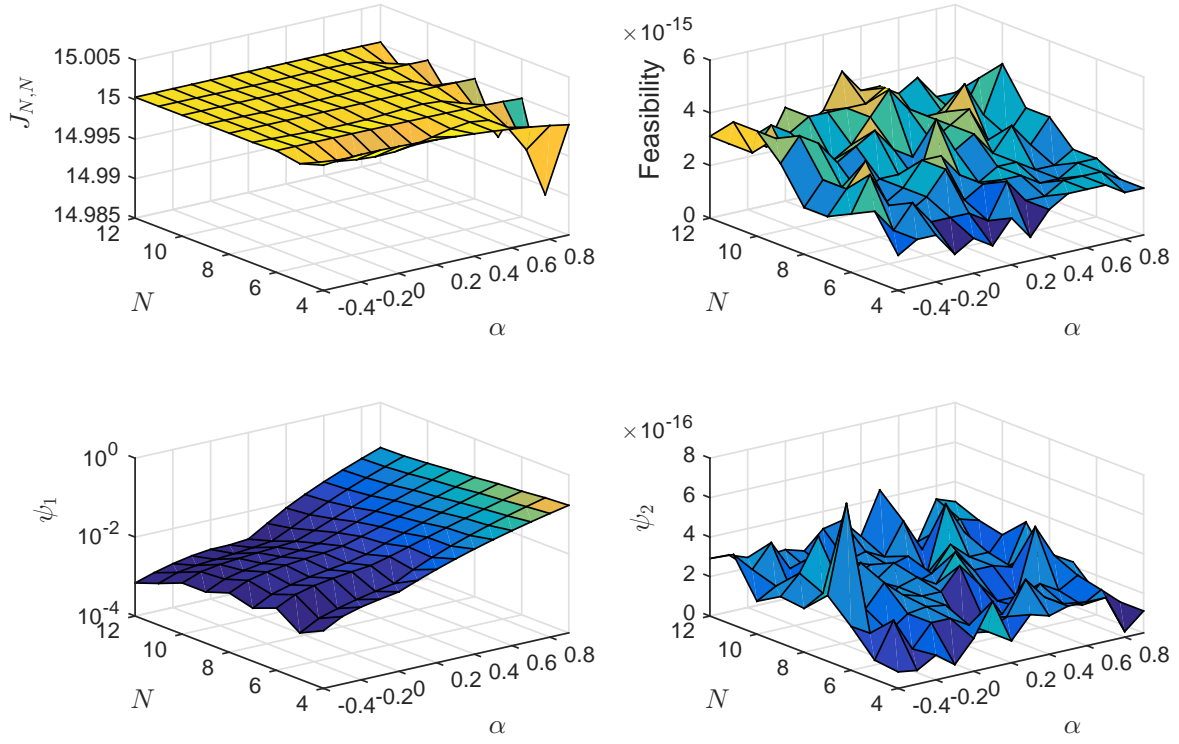


Figure 1. The figure shows the plots of the approximate optimal cost functional  $J_{N,N}^*$  (upper left), the feasibility of the optimal solution  $\mathbf{Z}^*$  as reported by the solver (upper right), the maximum error in the initial condition (2.3),  $\psi_1$ , at the 101 linearly spaced nodes  $(x_i, y_i)$  in the  $y$ - and  $t$ -directions from 0 to 4, and 0 to 1, respectively in semi-logarithmic scale (lower left), and the maximum error in the boundary condition (3.6),  $\psi_2$  (lower right). All of the plots were generated using the same 101 points  $(x_i, y_i)$ .

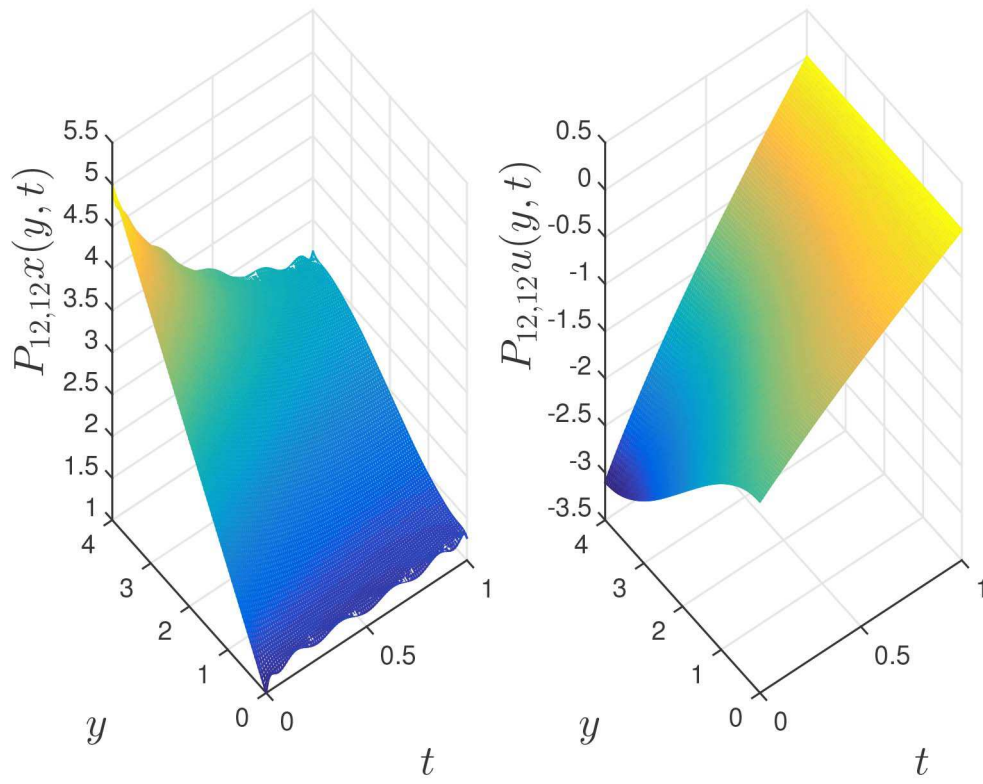


Figure 2. The figure shows the state and control profiles on  $D_{4,1}^2$  using  $N = 12$  and  $\alpha = -0.2$ .

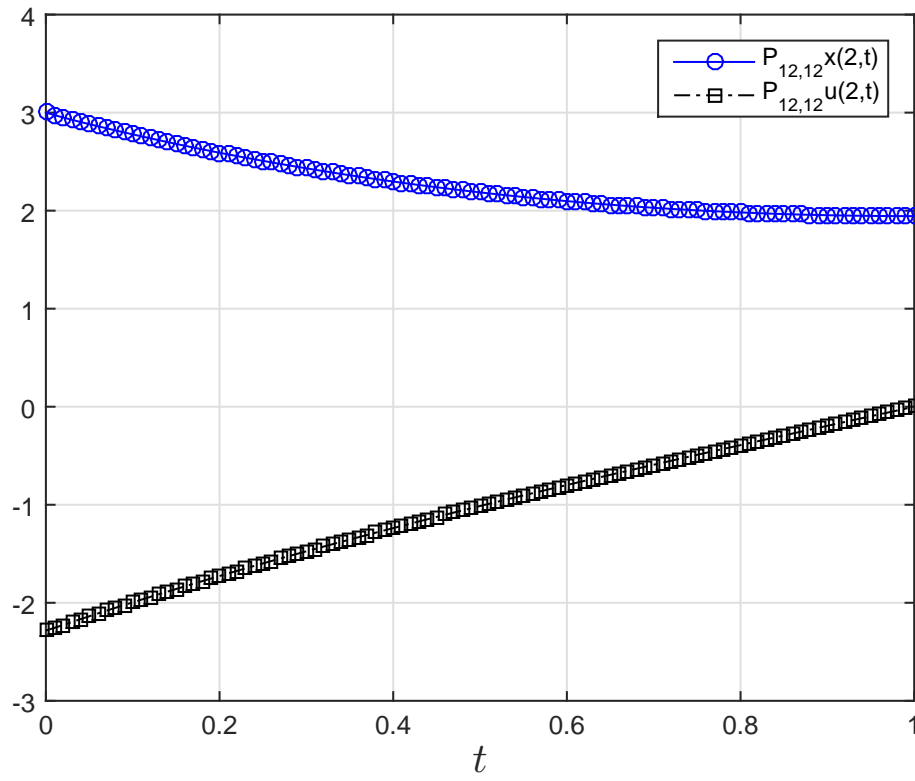


Figure 3. The figure shows the state and control profiles at the midpoint  $y = 2$  using  $N = 12$  and  $\alpha = -0.2$ .